(19)日本国特許庁（JP）　(12)公開特許公報（A）　(11)特許出願公開番号
特開2000－20501
（P2000－20501A）
(43)公開日　平成12年1月21日（2000.1.21）

(51)Int.Cl.'	識別記号	FI		テーマコード（参考）
G06F 17/16		G06F 15/31	Z	5B045
15/16	390	15/16	390Z	5B056

審査請求　未請求　請求項の数7　OL　（全17頁）

(21)出願番号　特願平10－188449

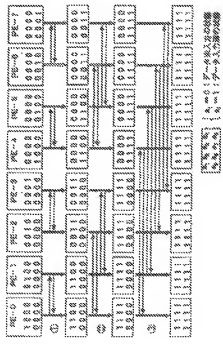
(22)出願日　平成10年7月3日（1998.7.3）

(71)出願人　000003078
株式会社東芝
神奈川県川崎市幸区堀川町72番地
(72)発明者　上松　善実
神奈川県横浜市磯子区新杉田町8番地　株式会社東芝横浜事業所内
(74)代理人　100083161
弁理士　外川　英明

Fターム（参考）　5B045 AA07 BB02 BB28 BB47 GG02 GG12
5B056 AA04 BB45 DD14 FF05

(54)【発明の名称】　並列計算機システム及びその演算処理装置間の通信方法

(57)【要約】
【課題】　並列計算機の各演算処理装置で分散処理されたデータを効率的に集結する。
【解決手段】　識別番号0，1，…，2^n－1が付与された2^n台の演算処理装置と識別記憶装置及び通信手段を備えた並列計算機システムで、2^n個の小配列に分割して各演算処理装置に分配／演算処理されたデータ配列を1つの配列に集約する際に、識別番号Nに対し2進法で表した識別番号Nの2^iの位の数を反転させた番号N'を対応させ、識別番号Nの演算処理装置と識別番号N'の演算処理装置の間でデータ配列の演算処理結果を相互に通信する操作iを$i=0$から$i=n-1$まで順次行う。この際、$j>i$となるjに対しては、操作iの際に、識別番号N，N'の演算処理装置間である演算処理装置による演算処理結果に加えて操作（j－i）までで得られた演算処理結果を通信する。



【特許請求の範囲】

【請求項1】　固有の識別子を有する少なくとも2^n台の演算処理装置と、これら各演算処理装置に各々対応する識別記憶装置および通信手段とを備え、この通信手段により各演算処理装置間でデータの授受を行う並列計算機システムにおいて、2^n個の小配列に分割して2^n台の演算処理装置に分配され各演算処理装置で演算処理されたデータ配列を得ず1つの配列に集結する際に、2^n台の演算処理装置に識別番号0，1，…，2^n-1を付与し，識別番号Nの演算処理装置に対し2進法で表した識別番号Nの2^iの位の数を反転させた番号N'を識別番号とする演算処理装置を対応させ、前記データ配列の演算処理結果と識別番号N'の演算処理装置の間で相互に送受信する操作を$i=0$から$i=n-1$まで順次行い，$j>0$なるjに対しては、操作jの際に、識別番号N，N'の演算処理装置間で各演算処理装置による演算処理結果に加えて操作$(j-1)$までで得られた演算処理結果を送受信することにより2^j台の演算処理装置間で各回の操作でデータ配列を集結させることを特徴とする並列計算機システム。

【請求項2】　固有の識別子を有する(2^n+k)台の演算処理装置と、これら各演算処理装置に各々対応する識別記憶装置および通信手段とを備え、この通信手段により各演算処理装置間でデータの授受を行う並列計算機システムにおいて、(2^n+k)個の小配列に分割して(2^n+k)台の演算処理装置に分配・演算処理されたデータ配列を再び1つの配列に集結する際に、前記(2^n+k)台の演算処理装置に識別番号0，1，…，2^n+k-1を付与し，識別番号Nの演算処理装置に対し2進法で表した識別番号Nの2^iの位の数を反転させた番号N'を識別番号とする演算処理装置を対応させ、前記データ配列の演算処理結果を識別番号Nの演算処理装置と識別番号N'の演算処理装置の間で相互に送受信する操作を$i=0$から$i=m$まで順次行い、$j>0$なるjに対しては、操作jの際に、$N\leq 2^n+k$なる識別番号Nの演算処理装置からはその演算処理装置の演算処理結果及び操作$(j-1)$までで得られた演算処理結果を送信し，$N>2^n+k$なる識別番号Nの演算処理装置からは操作$(j-1)$までで得られた演算処理結果を送信することにより(2^n+k)台の演算処理装置において$(m+1)$回の操作でデータ配列を集結させることを特徴とする並列計算機システム。

【請求項3】　固有の識別子を有する(2^n+k)台の演算処理装置と、これら各演算処理装置に各々対応する識別記憶装置および通信手段とを備え、この通信手段により各演算処理装置間でデータの授受を行う並列計算機システムにおいて、(2^n+k)個の小配列に分割して(2^n+k)台の演算処理装置に分配・演算処理されたデータ配列を再び1つの配列に集結する際に、前記(2^n+k)台の演算処理装置に識別番号0，1，…，2^n+k-1を付与し，識別番号Nの演算処理装置に対し2進法で表した識別番号Nの2^iの位の数を反転させた番号N'を識別番号とする演算処理装置を対応させ、前記データ配列の演算処理結果を識別番号Nの演算処理装置と識別番号N'の演算処理装置の間で相互に送受信する操作を$i=0$から$i=n-1$まで順次行い，$j>0$なるjに対しては、操作jの際に、識別番号N，N'の演算処理装置間で各演算処理装置による演算処理結果に加えて操作$(j-1)$までで得られた演算処理結果を送受信することにより(2^n+k)台の演算処理装置において$(m+1)$回の操作でデータ配列を集結させることを特徴とする並列計算機システム。

【請求項4】　$n>m$なるn，mについて、固有の識別子を有する(2^n+2^m)台の演算処理装置と、これら各演算処理装置に各々対応する識別記憶装置および通信手段とを備え、この通信手段により各演算処理装置間でデータの授受を行う並列計算機システムにおいて、(2^n+2^m)個の小配列に分割して(2^n+2^m)台の演算処理装置に分配・演算処理されたデータ配列を再び1つの配列に集結する際に、前記(2^n+2^m)台の演算処理装置を2つ分かちなるグループG_1と2つ分かちなるグループG_2に分割し，また前記データ配列を初めの2^n個の小配列からなる配列A_1とその後の2^m個の小配列からなる配列A_2の2つに分割し，この配列A_1，A_2をそれぞれグループG_1，G_2と対応づけて分配，演算処理を行い、グループG_1の2^n台の演算処理装置に識別番号0，1，…，2^n-1を付与し，識別番号Nの演算処理装置に対し2進法で表した識別番号Nの2^iの位の数を反転させた番号N'を識別番号とする演算処理装置を対応させ、前記データ配列の演算処理結果を識別番号Nの演算処理装置と識別番号N'の演算処理装置の間で相互に送受信する操作を$i=0$から$i=n-1$まで順次行い，$j>0$なるjに対して、操作jの際に識別番号N，N'の演算処理装置間で各演算処理装置による演算処理結果に加えて操作$(j-1)$までで得られた演算処理結果を送受信することにより2^j台の演算処理装置において$(n+1)$回の操作でデータ配列を集結させることを特徴とする並列計算機システム。

【請求項5】　$n>m$なるn，mについて、固有の識別子を有する(2^n+2^m)台の演算処理装置と、これら各演算処理装置に各々対応する識別記憶装置および通信手段とを備え、この通信手段により各演算処理装置間でデータの授受を行う並列計算機システムにおいて、(2^n+2^m)個の小配列に分割して(2^n+2^m)台の演算処理装置に分配・演算処理されたデータ配列を再び1つの配列に集結する際に、前記(2^n+2^m)台の演算処理装置を2つ分かちなるグループG_1と2つ分かちなるグループG_2に分割し，また前記データ配列を初めの2^n個の小配列からなる配列A_1とその後の2^m個の小配列からなる配列A_2の2つに分割し，この配列A_1，A_2をそれぞれグループG_1，G_2と対応づけて分配，演算処理を行い、グループG_1の2^n台の演算処理装置に識別番号0，1，…，2^n-1を付与し，識別番号Nの演算処理装置に対し2進法で表した識別番号Nの2^iの位の数を反転させた番号N'を識別番号とする演算処理装置を対応させ、前記データ配列の演算処理結果を識別番号Nの演算処理装置と識別番号N'の演算処理装置の間で相互に送受信する操作を$i=0$から$i=n-1$まで順次行い，$j>0$なるjに対しては、操作jの際に、識別番号N，N'の演算処理装置間で各演算処理装置による演算処理結果に加えて操作$(j-1)$までで得られた演算処理結果を送受信することにより2^j台の演算処理装置において$(n+1)$回の操作でデータ配列を集結させることを特徴とする並列計算機システム。

識番号ｎとする演算処理装置を対応させ、前記データ配列の演算処理結果を識別番号Ｎの演算処理装置と識別番号Ｎ’の演算処理装置の間で相互に送受信する操作ｊ＝ｉ＋１からｊ＝ｎ－１まで順次行い、ｊ＞０なる）に対して、操作ｊの際に識別番号Ｎ、Ｎ’の演算処理装置間で各演算処理装置による演算処理結果に加えて操作｛ｊ－１｝までで得られた演算処理結果を送受信することによりグループＧ０内で配列Ａ，を集結させる第２の工程と、グループＧ０からグループＧｎの各演算処理装置に配列Ａ，を、グループＧ０からグループＧｉの各演算処理装置に配列Ａ，を送信する第３の工程とを有し、第１の工程と第２の工程を並列に実行した後に第３の工程を行なうことにより｛２’＋２’｝台の演算処理装置においてデータ配列を集結させることを特徴とする並列計算機システム。

【数１】

$$2^{n_1}+2^{n_2}+2^{n_3}+\cdots+2^{n_k}$$

個の台配列（但し、$n_1>n_2>n_3>\cdots>n_k\geq 0$）に分割して

【数２】

$$2^{n_1}+2^{n_2}+2^{n_3}+\cdots+2^{n_k}$$

台の演算処理装置に分配・演算処理されたデータ配列を再び１つの配列に集結する際に、これらの演算処理装置のうち

【数３】

$$2^{n_1},\ 2^{n_2},\ 2^{n_3},\ \cdots,\ 2^{n_k}$$

台をそれぞれグループＧ１，Ｇ２，…，Ｇｋとしてｋ個のグループに分割するとともに、前記小配列のうち

【数４】

$$2^{n_1},\ 2^{n_2},\ 2^{n_3},\ \cdots,\ 2^{n_k}$$

個の小配列をそれぞれ配列Ａ１，Ａ２，…，Ａｋとしてｋ個の配列の分割し、このｋ個の配列とｋ個のグループＧ１，Ｇ２，…，Ｇｋとを１対１に対応づけて分配・演算処理を行い、１≦ｐ≦ｋなる各ｐに対し、グループＧｐの（そのｎ，乗）台の演算処理装置に識別番号０，１，…を付与し、識別番号Ｎの演算処理装置に対し２進法で表した識別番号Ｎの２’の位の数を反転させた番号Ｎ’を識別番号とする演算処理装置を対応させ、前記データ配列の演算処理結果を識別番号Ｎの演算処理装置と識別番号Ｎ’の演算処理装置の間で相互に送受信する操作ｉ＝０からｉ＝ｎ－１まで順次行い、ｊ＞０なるｊに対して、操作ｊの際に識別番号Ｎ，Ｎ’の演算処理装置間で各演算処理装置による演算処理結果に加えて操作｛ｊ－１｝までで得られた演算処理結果を送受信することによりグループＧ０内の演算処理装置でデータ配列Ａ，を集結させるグループ内工程ｐを実行し、グループ内工程｛ｋ－１｝が終了した後、グループＧ１の演算処理装置群から配列Ａ１の演算結果をグループＧ２，，の演算処理装置に送信するグループ間工程ａを実行し、次に、グループＧ２の各演算処理装置に集結された配列Ａ２の演算結果を、グループＧ３の演算処理装置からｐ＞２なる全てのｐに対しグループＧｐに属する各演算処理装置に送信するとともに、グループＧ２の演算処理装置群から、及びグループＧ３，，の演算処理装置から受信した配列Ａ３，，，Ａ４の演算結果をグループＧ４，，の演算処理装置に送信するグループ間工程ｐを、ｐ＝ｋ－１からｐ＝２までｐに関して降順に実行することにより、

【数５】

$$2^{n_1}+2^{n_2}+2^{n_3}+\cdots+2^{n_k}$$

台の演算処理装置においてデータ配列を集結させることを特徴とする並列計算機システム。

【請求項６】　請求項４または記載の並列計算機システムを用いて演算処理装置のグループ間でのデータ交換を行う場合、ｐ＞ｑなるｐ，ｑについて、２’台の演算処理装置からなるグループＧｐで集結され共有されているデータ配列Ａとと、２’台の演算処理装置からなるグループＧｑで集結され共有されているデータ配列Ｂとを、グループＧｐ，Ｇｑ間で相互に送受信する際に、グループＧｐのなかから選択される２’台の演算処理装置をグループＧｐの各演算処理装置と１対１に対応させてグループＧｑの各演算処理装置にデータ配列Ａを送信する操作を並列に実施するとともに、グループＧｑも、それぞれが２’’’台の演算処理装置からなる小グループα，α，…，α，｛ｒ＝２’｝に分割して、各々のグループＧｑのｒ台の各演算処理装置と１対１に対応させ、小グループαのなかから選択される１台の演算処理装置に対して、小グループαに対応するグループＧｑの演算処理装置からデータ配列Ｂを送信した後、小グループαの演算処理装置間でデータ配列Ｂを送受信する操作ｉを、１≦ｉ≦ｓなるｉに関して並列に実行するすることにより、２’台の演算処理装置と２’台の演算処理装置にデータ配列Ａとデータ配列Ｂを共有させることを特徴とする並列計算機システム。

【請求項７】　請求項１乃至６のいずれか記載の並列計算機システムを用いて２台の演算処理装置間でデータを交換する工程は、演算処理装置の識別番号の大きい方から小さい方へデータを送る第１の送信工程と、演算処理装置の識別番号の小さい方から大きい方へデータを送る第２の送信工程とからなり、この第１の送信工程と第２の送信工程のうちから選択される１工程を先に行った後、続いて他の１工程を行うことを特徴とする並列演算機システムの演算処理装置間の通信方法。

【発明の詳細な説明】
【０００１】
【発明の属する技術分野】本発明は，通信手段および個別記憶装置を備えた多数の演算処理装置からなり，特に並列計算を目的とした並列計算機システム及びその演算処理装置の通信方法に関する。
【０００２】
【従来の技術】原子力施設をはじめとする大規模な施設の設計においては，例えば遮蔽設計などにおける放射線挙動計算，炉心設計における炉心核熱水力解析などの大規模な計算がかなりの頻度で要求される。この要求に応えるためには大幅な計算速度の向上が必要である。このため最近では，通信手段と個別の記憶装置を備えた多数の演算処理装置を用いて，１台の演算処理装置しか持たない計算機を使用していたのでは得られないような高速度で，解析を行うことが考案されている。
【０００３】例えば炉心設計であれば，原子炉の炉心を複数の燃料集合体からなる幾つかのセグメントに分割し，それぞれのセグメントを１つの演算処理装置に対応させて，出力計算と熱水力計算を各々の演算処理装置で並列に計算させる。セグメント間での中性子束の流出入およびチャンネル間の冷却材の圧力バランスを解析する際には，前記通信手段によりセグメント境界の中性子＊束，各チャンネルの圧力損失のデータを演算処理装置間でやり取りすることで，空間的に連続した解析が行われる。
【０００４】また，遮蔽設計であれば，例えば原子炉の炉心，冷却材，遮蔽体などを含む全体系を幾つかの小領域に分割し，それぞれの小領域を１つの演算処理装置に対応させて，放射線束分布計算を各々の演算処理装置で並列に計算させる。小領域間での中性子束の流出入を解析する際には，前記通信手段により小領域境界の中性子束のデータを演算処理装置間でやり取りすることで，空間的に連続した解析が行われる。
【０００５】
【発明が解決しようとする課題】複数の演算処理装置を用いて並列に計算を行わせる際に，演算処理装置間の通信を行うことなく全く独立に計算を進めることができる例はまれであり，通常は演算処理装置間の通信を行いながら計算を進める。たとえば，４行４列の行列A，Bの掛け算を４台の演算処理装置で実施して４行４列の行列Cを求める場合を考える。A，B，Cの要素をそれぞれa_{ij}，b_{ij}，c_{ij}で以下のように表記する。
【０００６】
【数６】

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \times \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \\ b_{41} & b_{42} & b_{43} & b_{44} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \\ c_{31} & c_{32} & c_{33} & c_{34} \\ c_{41} & c_{42} & c_{43} & c_{44} \end{bmatrix}$$

このとき，４台の演算処理装置のうちの１台においては例えば，
【０００７】
【数７】

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \end{bmatrix} \times \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \\ b_{41} & b_{42} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

のように計算が行われる。
【０００８】この例から明らかなように，演算に使う側（a_{ij}またはb_{ij}）については行或いは列全体について の要素のデータが必要である。また，演算の結果として得られるc_{ij}の方は，各々の演算処理装置に於いては部分的にしかデータが得られない。このことは，例えば次のステップで行列Cと行列Aの掛け算を行う必要が生じたとき，計算で得られた要素だけではデータに不足が生じることを意味する。したがって，A×B＝Cの計算を実施した後で残りの部分，上の式で言えば行列Cの少なくとも第１行と第２行のデータ及び第１列と第２列のデータは満たされた状態にしておかねばならない。
【０００９】これらの問題を一般化すると次のようになる。（n×k）個からなる配列X(nk)があり，これがn台の演算処理装置に分割され，例えば識別番号１の演算処理装置ではX(1)，X(2)，…X(k)，識別番号２の演算処理装置ではX(k+1)，X(k+2)，…，X(2k)の計算結果を持っているものとする。この状態からn台の演算処理装置の間で通信を行うことにより，n台の演算処理装置が配列X(nk)の計算結果を持っている状況を作る操作が必要となることがある。
【００１０】このときの通信は「対」であることが通信手段上の条件である。すなわち，例えば演算処理装置１から演算処理装置２にデータを転送する際には，演算処理装置２は演算処理装置１からデータを受けとる態勢になければならないのであって，このとき演算処理装置２が他の処理，例えば演算処理装置３にデータを転送しようとしたり演算処理装置４からデータを受けようとしたりすると，通信は失敗して計算は中断することとなる。通信が滞りなく行われるには送信側と受信側の混乱がないように通信の順序を予め決めておく必要がある。
【００１１】４台の演算処理装置を使う場合を例にとれば，容易に考えられる方法として次のものが挙げられ

る。以下、表記を簡略化するため演算処理装置１，２，＊　＊３，４をそれぞれ＃１，＃２，＃３，＃４と書く。

（１）　送信…受信を１つずつ順次行う方法

[1] ＃１の計算結果→＃２．
[2] ＃３の計算結果→＃１，
[3] ＃１の計算結果→＃４，
[4] ＃２の計算結果→＃１，
[5] ＃２の計算結果→＃３，
[6] ＃４の計算結果→＃２，
[7] ＃３の計算結果→＃１，
[8] ＃３の計算結果→＃４，
[9] ＃４の計算結果→＃２，
[10] ＃４の計算結果→＃３

を順次実行する。

【００１２】ここで、[1]，[2]，[3]，…は処理のステップの番号を示す。演算処理装置をＮ台，１台に割り当てられたデータ量をｗとすれば，通信回数は

$2\times {}_{N}C_{1}=N(N-1)$

であり，データ移動量は

$2w\times {}_{N}C_{2}=wN(N-1)$

・データ集結

[1] ＃２の計算結果→＃１，　[2] ＃３の計算結果→＃１，
『1] ＃４の計算結果→＃１、…　を順次実行。
＃１に全データが揃う。

・全データ配布

[1] ＃１→＃２．　[2] ＃１→＃３，　[3] ＃１→＃４．…

を順次実行。　配列全体を＃２，＃３，＃４に送信する。

【００１４】この場合の通信回数は　２（Ｎ－１）回，データ移動量は集結時に（Ｎ－１）ｗ，配布時に　Ｎ（Ｎ－１）ｗである。この方法は（１）の方法に比べて通信回数は少ないが，全データ配布時に送信されるデー★

[1] ＃１の計算結果→＃２，
[2] ＃１の計算結果→＃３，
[3] ＃１の計算結果→＃４，
[4] ＃２の計算結果→＃１，
[5] ＃３の計算結果→＃１，
[6] ＃４の計算結果→＃１，

【００１５】この通信方法によれば，通信が成長することも衰突することもなく，全データが４台の演算処理装置に行き渡る。演算処理装置が４台であれば通信回数は２（Ｎ－１），データ移動量は２（Ｎ－１）ｗである。Ｎ＝４であれば通信に要する時間は前述の（１）の方法の半分である。Ｎが大きくなるとともに差は広がる。

【００１７】（４）　演算処理装置の Binary tree により代表の演算処理装置にデータを集めた後，各演算処理装置に配布する。これは（２）の方法を改良したもので、例えば次のように行う。

・データ集結

[1] ＃２の計算結果…＃１，＃４の計算結果→＃３を同時に実行。
[2] ＃３に集結された計算結果→＃１
・全データ配布
[3] ＃１→＃３
[4] ＃１→＃２，　＃３→＃４を同時に実行。

[2] ＃１の計算結果→＃３，
[4] ＃２の計算結果→＃１，
[5] ＃２の計算結果→＃４，
[6] ＃３の計算結果→＃２，
[10] ＃４の計算結果→＃３

※である。Ｎ＝４ならば通信回数は上述の12回である。この方法によれば，時間はかかるが通信上の混乱は避けられる。なお，${}_{p}C_{q}$ はｐ個の要素からｑ個の要素を選ぶ組合せの数を示す。

【００１３】（２）　代表の演算処理装置にデータを集めた後，各演算処理装置に配布する。

★タ量が多い点が短所である。

【００１５】また，通信の効率化を図った手法として次のものがある。

（３）　演算処理装置の［対］の組み合わせに対して並列・隣接的に通信を行う。　これは（１）の方法を改良したもので，例えば次のように行う。

＃３の計算結果→＃４　を同時に実行。
＃２の計算結果→＃４　を同時に実行。
＃２の計算結果→＃３　を同時に実行。
＃４の計算結果→＃３　を同時に実行。
＃４の計算結果→＃２　を同時に実行。
＃３の計算結果→＃２　を同時に実行。

【００１８】この方法化，よれば，演算処理装置がＮ台であれば，通信回数は　２×log Ｎ回，データ通信量は，集結時に（Ｎ－１）ｗ，配布時に　Ｎｗ log Ｎである。Ｎ＝４であれば通信回数は（２）の方法の２／３，データ移動量は（２）の方法の　11／15である。Ｎが大きくなるとともに差は広がる。

【００１９】（３）の方法は（４）の方法に比べてデータ移動量は少ないが通信回数が多いため，扱う配列が小さい場合には適していない。（４）の方法は通信回数は少ないが，データ移動量が多いため，巨大な配列を扱う場合には適していない。

【００２０】よって，データ移動量と通信回数がともに最適化されたい，あらゆる条件に対して適用可能な一般化された手法が必要である。本発明は、このような点を考慮してなされたもので、通信によるデータの授受を並列化行えるようにすることで，演算処理装置間の通信回数およびデータの授受の際の待ち時間を最小限に抑えて高

速化を図ることができる並列計算機システム及びその演算処理装置間の通信方法を提供することを目的とする。

【００２１】

【課題を解決するための手段】上記目的を達成するため、本発明の請求項１記載の発明は、固有の識別子を有する少なくとも２^ｎ台の演算処理装置と、これら各演算処理装置と各々対応する識別記憶装置および通信手段とを備え、この通信手段により各演算処理装置間でデータの授受を行う並列計算機システムにおいて、２^ｎ個の小配列に分割して２^ｎ台の演算処理装置に分配され演算処理されたデータ配列を再び１つの配列に集結する際に、２^ｎ台の演算処理装置に識別番号０，１，…，２^ｎ－１を付与し、識別番号Ｎの演算処理装置に対し２進法で表した識別番号Ｎの２^ｎの位の数を反転させた番号Ｎ’を識別番号とする演算処理装置を対応させ、前記データ配列の演算処理結果を識別番号Ｎの演算処理装置と識別番号Ｎ’の演算処理装置の間で相互に送受信する操作ｉをｉ＝０からｉ＝ｎ－１まで順次行い、ｊ＞０なるｊに対しては、操作ｉの際に、識別番号Ｎ、Ｎ’の演算処理装置間で各演算処理装置による演算処理結果に加えて操作（ｊ－１）までで得られた演算処理結果を送受信することにより２^ｎ台の演算処理装置間でｍ回の操作でデータ配列を集結させることを特徴とする。

【００２２】また、請求項２記載の発明は、固有の識別子を有する（２^ｎ＋ｋ）台の演算処理装置と、これら各演算処理装置に各々対応する識別記憶装置および通信手段とを備え、この通信手段により各演算処理装置間でデータの授受を行う並列計算機システムにおいて、（２^ｎ＋ｋ）個の小配列に分割して（２^ｎ＋ｋ）台の演算処理装置に分配・演算処理されたデータ配列を再び１つの配列に集結する際に、前記（２^ｎ＋ｋ）台の演算処理装置に識別記憶手段及び通信手段を備えた（２^ｎ－ｋ）台の演算処理装置群を形成し、この演算処理装置群を構成する２^ｎ台の演算処理装置に識別番号０，１，…，２^ｎ－１を付与し、識別番号Ｎの演算処理装置に対し２進法で表した識別番号Ｎの２^ｎの位の数を反転させた番号Ｎ’を識別番号とする演算処理装置と前記（２^ｎ＋ｋ）台の演算処理装置群を構成する２^ｎ台の演算処理装置群とを対応させ、前記データ配列の演算処理結果を識別番号Ｎの演算処理装置と識別番号Ｎ’の演算処理装置の間で相互に送受信する操作ｉをｉ＝０からｉ＝ｍ－１まで順次行い、ｊ＞０なるｊに対しては、操作ｉの際に、識別番号Ｎ、Ｎ’の演算処理装置間で各演算処理装置による演算処理結果に加えて操作（ｊ－１）までで得られた演算処理結果を送受信することにより２^ｎ台の演算処理装置間でｍ回の操作でデータ配列を集結させることを特徴とする。

【００２３】また、請求項３記載の発明は、固有の識別子を有する（２^ｎ＋ｋ）台の演算処理装置と、これら各演算処理装置に各々対応する識別記憶装置および通信手段とを備え、この通信手段により各演算処理装置間でデータの授受を行う並列計算機システムにおいて、（２^ｎ＋ｋ）個の小配列に分割して（２^ｎ＋ｋ）台の演算処理装置に分配・演算処理されたデータ配列を再び１つの配列に集結する際に、前記（２^ｎ＋ｋ）台の演算処理装置に識別記憶手段及び通信手段を備えた（２^ｎ－ｋ）台の演算処理装置群を形成し、この演算処理装置群を構成する２^ｎ台の演算処理装置に識別番号０，１，…，２^ｎ－１を付与し、識別番号Ｎの演算処理装置に対し２進法で表した識別番号Ｎの２^ｎの位の数を反転させた番号Ｎ’を識別番号とする演算処理装置と前記（２^ｎ＋ｋ）台の演算処理装置群を構成する２^ｎ台の演算処理装置群とを対応させ、前記データ配列の演算処理結果を識別番号Ｎの演算処理装置と識別番号Ｎ’の演算処理装置の間で相互に送受信する操作ｉをｉ＝０からｉ＝ｎ－１まで順次行い、ｊ＞０なるｊに対しては、操作ｉの際に、識別番号Ｎ、Ｎ’の演算処理装置間で各演算処理装置による演算処理結果に加えて操作（ｊ－１）までで得られた演算処理結果を送受信することにより２^ｎ台の演算処理装置間でｍ回の操作でデータ配列を集結させることを特徴とする。

【００２４】また、請求項４記載の発明は、ｎ＝ｍなるｎ，ｍについて、固有の識別子を有する（２^ｎ＋２^ｎ）台の演算処理装置と、これら各演算処理装置に各々対応する識別記憶装置および通信手段とを備え、この通信手段により各演算処理装置間でデータの授受を行う並列計算機システムにおいて、（２^ｎ＋２^ｎ）個の小配列に分割して（２^ｎ＋２^ｎ）台の演算処理装置に分配・演算処理されたデータ配列を再び１つの配列に集結する際に、前記（２^ｎ＋２^ｎ）台の演算処理装置を２^ｎ台からなるグループＧ，と２^ｎ台からなるグループＧ２に区分割し、また前記データ配列を初めの２^ｎ個の小配列からなる配列Ａ，とその後の２^ｎ個の小配列からなる配列Ａ２の２つに分割し、この配列Ａ，、Ａ２をそれぞれグループＧ，、Ｇ２と対応づけて分配、演算処理を行い、グループＧ，の２^ｎ台の演算処理装置に識別番号０，１，…，２^ｎ－１を付与し、識別番号Ｎの演算処理装置に対し２進法で表した識別番号Ｎの２^ｎの位の数を反転させた番号Ｎ’を識別番号とする演算処理装置を対応させ、前記データ配列の演算処理結果を識別番号Ｎの演算処理装置と識別番号Ｎ’の演算処理装置の間で相互に送受信する操作ｉをｉ＝０からｉ＝ｎ－１まで順次行い、ｊ＞０なるｊに対して、操作ｊの際に識別番号Ｎ，Ｎ’の演算処理装置間で各演算処理装置による演算処理結果に加えて操作（ｊ－１）までで得られた演算処理結果を送受信することによりグループＧ，内でデータ配列を集結させる第１の工程と、グループＧ２の２^ｎ台の演算処理装置に識

別番号0，1，…，2^n-1を付与し、識別番号Nの演算処理装置に対し2進法で表した識別番号Nの2^jの位の数を反転させた番号N'を識別番号とする演算処理装置を対応させ，前記データ配列の演算処理結果を識別番号Nの演算処理装置と識別番号N'の演算処理装置の間で相互に送受信する操作jをj＝0からj＝n－1まで順次行い，j＞0なるjに対して，操作jの際に識別番号N，N'の演算処理装置間で各演算処理装置による演算処理結果に加えて操作（j－1）までで得られた演算処理結果を送受信することによりグループG_1内でデータ配列を集結させる第2の工程と，グループG_1からグループG_2の各演算処理装置に配列A_1を，グループG_2からグループG_1の各演算処理装置に配列A_2を送信する第3の工程とを有し，第1の工程と第2の工程を並列に実行した後に第3の工程を行なうことにより（2^m+2^n）台の演算処理装置においてデータ配列を集結させることを特徴とする。

【0025】また，請求項5記載の発明は，固有の識別子を有する複数の演算処理装置と，これら各演算処理装置に各々対応する個別記憶装置および通信手段とを備えた並列計算機システムにおいて，

【数8】

$$2^{n_1}+2^{n_2}+2^{n_3}+\cdots+2^{n_k}$$

個の小配列（但し，$n_1>n_2>n_3>\cdots>n_k\geq 0$）に分割して

【数9】

$$2^{n_1}+2^{n_2}+2^{n_3}+\cdots+2^{n_k}$$

台の演算処理装置に分配・演算処理されたデータ配列を再び1つの配列に集結する際に，これらの演算処理装置のうち

【数10】

$$2^{n_1},2^{n_2},2^{n_3},\cdots,2^{n_k}$$

台をそれぞれグループG_1，G_2，…，G_kとしてk個のグループに分割するとともに，前記小配列のうち

【0026】

【数11】

$$2^{n_1},2^{n_2},2^{n_3},\cdots,2^{n_k}$$

個の小配列をそれぞれ配列A_1，A_2，…，A_kとしてk個の配列に分割し，このk個の配列とk個のグループG_1，G_2，…，G_kとを1対1に対応づけて分配，演算処理を行い，1≦p≦kなる各pに対し，グループG_pの（2のn_p乗）台の演算処理装置に識別番号0，1，…を付与し，識別番号Nの演算処理装置に対し2進法で表した識別番号Nの2^jの位の数を反転させた番号N'を識別番号とする演算処理装置を対応させ，前記データ配列の演算処理結果を識別番号Nの演算処理装置と識別番号N'の演算処理装置の間で相互に送受信する操作jをj＝0からj＝n_p－1まで順次行い，j＞0なるjに対して，操作jの際に識別番号N，N'の演算処理装置間で各演算処理装置による演算処理結果に加えて操作（j－1）までで得られた演算処理結果を送受信することによりグループG_p内の演算処理装置でデータ配列A_pを集結させるグループ内工程pを実行し，グループ内工程（k－1）が終了した後，グループG_kの演算処理装置から配列A_kの演算結果をグループG_{k-1}の演算処理装置に送信するグループ間工程kを実行し，次に，グループG_pの各演算処理装置に集結された配列A_pの演算結果を，グループG_pの演算処理装置からq＞pなる全てのqに対しグループG_qに属する各演算処理装置に送信するとともに，グループG_pの演算処理装置から，グループG_p自身の演算結果である配列A_p及びグループG_{p+1}の演算処理装置から受信した配列A_{p+1}，…，A_kの演算結果をグループG_{p-1}の演算処理装置に送信するグループ間工程pを，p＝k－1からp＝2までpに関して降順に実行することにより，

【0027】

【数12】

$$2^{n_1}+2^{n_2}+2^{n_3}+\cdots+2^{n_k}$$

台の演算処理装置においてデータ配列を集結させることを特徴とする。

【0028】なお，この際には，k個のグループ内工程1，2，…，kを並列に実行し，1≦s≦k－1なるsに対して，グループ内工程sが終了した時点で順次グループ間工程（s＋1）を実行することで，全体の通信に要する時間をさらに短縮することができる。

【0029】また，請求項6記載の発明は，請求項4または5記載の並列計算機システムを用いて演算処理装置のグループ間でのデータ交換を行う場合，2^p台の演算処理装置からなるグループG_pで集結され共有されているデータ配列Aと，2^q台の演算処理装置（p＞q）からなるグループG_qで集結され共有されているデータ配列Bとを，グループG_p，G_q間で相互に送受信する際に，グループG_pのなかから選択される2^q台の演算処理装置をグループG_qの各演算処理装置と1対1に対応させてグループG_qの各演算処理装置にデータ配列Aを送信する操作を並列に実施するとともに，グループG_pを，それぞれが2^{p-q}台の演算処理装置からなる小グループα_1，α_2，…，α_r（$r=2^q$）に分割して，各々の小グループとグループG_qのr個の各演算処理装置とを1対1に対応させ，小グループα_iのなかから選択される1台の演算処理装置に対して，小グループα_iに対応するグループG_qの演算処理装置からデータ配列Bを送信した後，小グループα_iの演算処理装置間でデータ配列Bを送受信する操作iを，1≦i≦rなるiに関して並列に実行するするすることにより，2^p台の演算処理装置

装置と2^i台の演算処理装置にデータ配列Aとデータ配列Bを共有させることを特徴とする。

【0030】また、請求項7記載の発明は、請求項1乃至5のいずれか記載の並列計算機システムを用いて2台の演算処理装置間でデータを交換する工程は、演算処理装置の識別番号の大きい方から小さい方にデータを送る第1の送信工程と、演算処理装置の識別番号の小さい方から大きい方にデータを送る第2の送信工程とからなり、この第1の送信工程と第2の送信工程のうちから選択される1工程を先に行った後、続いて他の1工程を行うことを特徴とする。

【0031】

【発明の実施の形態】本発明の実施の形態について、以下、図面を参照して説明する。図1は並列計算機システムの構成例を示すブロック図である。ここに示した並列計算機システムは、1台のホストの計算機1と8台の演算処理装置2－1、2－2、…、2－8で構成されている。ホストの計算機1には記憶装置3と通信手段4、演算処理装置2－1、2－2、…、2－8の各々には、記憶装置5－1、5－2、…、5－8と通信手段6－1、6－2、…、6－8が備えられている。例えば、ホストの計算機で読み込んだ入力データ等は、通信手段4から通信手段6－1、6－2、…、6－8を通じて全演算処理装置に送信される。演算処理装置2－1、2－2、…、2－8では各々割り当てられた領域の計算を行い、必要に応じて演算処理装置間の通信によりデータの授受を行う。

【0032】図1に示した並列計算機システムの構成に基き、本発明にかかる並列計算機システムの第1の実施の形態について説明する。図2は本実施の形態における並列計算機システムの演算処理装置間の通信方法を時系列で示すチャートである。

【0033】演算処理装置2－1、2－2、…、2－8の識別番号をそれぞれ0、1、…、7とし、これらを2進法の3桁の数として表示するとそれぞれ(000, 001, 010,011, 100, 101, 110, 111)となる。8×n個のデータからなる配列Aがn個のデータからなる8個の小配列a_1、a_2、…、a_8に分割されて、8台の演算処理装置2－1、2－2、…、2－8に割り当てられている。それぞれの演算処理装置で割り当てられた小配列のデータに関する演算処理を行った後、配列Aの要素を全ての演算処理装置に於いて集めることを考える。なお、図2においてS(0ならば1、1ならば0とする)または1はそれぞれ分割された小配列を示しており、Uは計算結果が未入力の状態を、1は計算結果が入力済みの状態を示す。

【0034】第1ステップとして、2^i の位の数を反転(0ならば1、1ならば0とする)させた数を識別番号としても演算処理装置との間でデータを交換する。例えば演算処理装置0(000)は演算処理装置4(100)、演算処理装置1(001)は演算処理装置5(101)、演算処理装置2(010)とn個のデータを交換する。各演算処理装置にn個の要素が集まる。

【0035】第2ステップでは、2^i の位の数を反転させた数を識別番号としても演算処理装置との間でデータを交換する。例えば演算処理装置0(000)は演算処理装置2(010)、演算処理装置3(011)は演算処理装置1(001)とデータを交換する。この時、例えば演算処理装置0から演算処理装置2への通信では、演算処理装置0自身による演算結果の他に第1ステップで演算処理装置1から受信したデータを含む2n個のデータを送信する。これにより各演算処理装置に4n個の要素が集まる。

【0036】最後に第3ステップとして、2^i の位の数を反転させた数を識別番号としても演算処理装置との間でデータを交換する。例えば演算処理装置0(000)は演算処理装置4(100)、演算処理装置3(011)は演算処理装置7(111)とn個のデータを交換する。各演算処理装置に8n個の要素が集まり、操作が完了する。

【0037】以上述べた通信方法は演算を2^i＝8個に分割した場合の3ステップの例である。同様に、演算を2^i＝N個に分割した台の演算処理装置において通信を行なう場合には、上述した各段の場合に比べてさらに1ステップが必要となり、全部でiステップとなる。一般に、演算をN個に分割し台の演算処理装置において通信を行う場合は、上述の方法を適用して、ステップ数 log$_2$N で通信が完了する。

【0038】本実施の形態の作用効果について以下概説する。例えば配列のデータ1つの大きさをM[word]、演算処理装置の台数をKとし、配列全体がK分割されて各演算処理装置に置かれているものとする。Kの値としては通例計算で最も一般的な条件である2のべき乗の場合、つまりK＝2^i と表される場合について考える。この状態から、演算処理装置間の通信によって演算処理装置全部が配列全体についてデータを把握している状態を作り出すのにかかる時間について考察する。一般にデータを通信するのに要する時間Tは

$$T = A + B \times M \qquad \cdots (1)$$

と表せる。ここで、Aは通信準備に要する時間で、通信するデータ量に関わらず1回の通信に必ず必要となる時間である。Aの値はデータ量に依らない。B×Mがデータ量に比例する項であり、Mがデータ量(word数)、Bが1word当たりの転送時間である。

【0039】データの授受のステップ数は log$_2$K＝n である。各ステップで演算処理装置等に送信と受信が1回づつ行われる。第iステップで授受されるデータ量は(M／K)×2^(i-1) [word]　である。データ量M[word]のデータを全演算処理装置において集結させるのに必要な送受信の回数は各演算処理装置当り2^i回であり、通受信する総データ量は

【0040】

15 16

[数18]

$$\sum_{m=1}^{n}\frac{M}{K}2^{m}=\frac{M}{K}2(2^{n}-1)=2M\left(1-\frac{1}{K}\right)\ [\mathrm{word}]$$

*である。よって、本発明を適用した場合の全通信時間Tは、

$$T(K)=2A\log_K K+2M\left(1-1/K\right)B \quad \cdots\cdots\cdots (6)$$

となる。

【0041】比較のため、従来法、例えばbinary treeの方式で1台の代表演算処理装置に全データを集めてあき、同様にbinary treeの方式で全演算処理装置にデータを配信する場合の通信時間を次に求めてみる。全データをK台の演算処理装置に集めるのに要する通信の回数は、代表演算処理装置においてn=$\log_2 K$回である。また、第jステップ（$n\geq j$）で送信されるデータ量は$(M/K)\times 2^{j-1}$[word]である。よって、代表演算処理装置に全データを集めるのにかかる時間T$_1$は、

$$T_1(K)=A\log_2 K+M\left(1-1/K\right)B \quad \cdots\cdots\cdots (7)$$

となる。

【0042】代表演算処理装置から各演算処理装置にデータを配布するj番のステップ数は$\log_2 K$で、演算処理装置あたり着信回数も最大で$\log_2 K$回である。ただし、各ステップ毎にM[word]のデータが送信される。よって、各演算処理装置にデータを配布する時にかかる時間T$_2$は、

$$T_2(K)=A\log_2 K+(M\log_2 K)B \quad \cdots\cdots\cdots (8)$$

となる。したがって、全通信時間Tは、$T=T_1+T_2$より

$$T(K)=2A\log_2 K+M\left(1-1/K+\log_2\right)B \quad \cdots\cdots (9)$$

となる。

【0043】図3及び図4のグラフは、横軸に演算処理装置台数、縦軸に通信に要する時間をとって、演算処理装置台数増加に伴う通信時間の増加の関係を示してある。従来のbinary treeの通信方式による(5)式の関係と、本実施の形態により通信を効率化した(2)式の関係を、比較して示している。このグラフ中の曲線のうち実線で示した符号10a、10bが本実施の形態の(2)式の場合、破線で示した符号11a、11bが従来の(5)式の場合を示している。図3に示した符号10a、11aを付した曲線は、通信されるデータ量が少なく、(1)式のA（通信立ち上げ時間）が全通信時間Tのほぼ半分を占める状況を、また図4に示した符号10b、11bを付した曲線は、通信されるデータ量が多く、(1)式のA（通信立ち上げ時間）が全通信時間Tに比べて十分小さい状況を想定している。このグラフからも明らかなように、本実施の形態によれば、演算処理装置の台数が少数の場合、多数の場合何れも従来の方法より通信に要する時間を少なくすることができる、すなわち、本実施の形態により、データの授受の際の待ち時間を最小限に抑え、計算の高速化を図ることができる。

【0044】なお、例えば演算処理装置の台数が16台からなる並列計算機システムにおいて、その内の8台の演算処理装置の間で上述の3ステップからなる配列の分割分配、演算結果の行うなど、複数の演算処理装置のうち2の冪乗の台数だけ抽出してこれら通信制御用の識別番号を付与し、この台数に適応して上述した方法で配列の分割分配、演算処理を行なうものとしてもよい。

【0045】上記第1の実施の形態においては、関係する演算処理装置の台数が2の冪乗であることを前提としている、一般的な条件として演算処理装置の台数が2の冪乗でない場合、すなわち台数が$2^{n}+k$等として表されない場合にも拡張したのが以下詳述する第2の実施の形態である。

【0046】本発明にかかる並列計算機システムの第2の実施の形態について説明する。ここでは、例えば並列計算の配列を8分割して、5台の演算処理装置（識別番号を0、1、…、5とする。）に割り当てる場合について説明する。図5は本実施の形態における並列計算機システムの演算処理装置間の通信方法を時系列で示すチャートである。この際のデータ処理には、図8台の演算処理装置のほかに2台の演算処理装置（識別番号を6，7とする。）を用いることとする。

【0047】第1ステップとして、2^{n}の位の数を反転させた数を識別番号としてもつ演算処理装置との間でデータを交換する。例えば演算処理装置0（000）は演算処理装置1（001）と、演算処理装置3（011）は演算処理装置2（010）と、それぞれ1個のデータを交換する。演算処理装置6（110）と演算処理装置7（111）は交換すべきデータがないので休止する、この時点で、演算処理装置0～5に2^{n}個のデータが集められる、

【0048】第2ステップでは、2^{1}の位の数を反転させた数を識別番号としてもつ演算処理装置との間でデータを交換する。例えば演算処理装置4（100）は演算処理装置6（110）とのデータ交換となる。この時点で演算処理装置6（110）は通信すべきデータがないので、演算処理装置4からデータを受信するのみとする。このデータ交換により、演算処理装置0～3および4番目のデータが、演算処理装置4～7には2^{n}個のデータが集められる、

【0049】第3ステップでは、2^{0}の位の数を反転させた数を識別番号としてもつ演算処理装置との間でデータを交換する。例えば演算処理装置0（000）は演算処理装置2（010）との交換である。演算処理装置0から演算

処理装置2へは2ｎ個のデータ、演算処理装置2から演算処理装置6へは4ｎ個のデータを送信する。このようにして6ｎ個のデータが8台の演算処理装置全てに行き渡る。

【0050】本実施の形態においては、一般に（2^ｋ）台の演算処理装置に対して、（2^ｍ－ｋ）台の演算処理装置を加えた2^ｍ台の演算処理装置群を構成し、この演算処理装置群に対して上述の第1の実施の形態で説明したステップにより並列計算を行うものとする。これにより、2の冪乗ではない台数の演算処理装置に対しても2の冪乗の場合に準じた構成とすることで、上記第1の実施形態と同様の作用効果を得ることができる。

【0051】次に本発明にかかる並列計算機システムの第3の実施の形態を説明する。本実施の形態における演算処理装置間の通信方法について、例として、配列を8個の小配列分割して8台の演算処理装置（識別番号0，1，…，5）に割り当てている場合について説明する。まず前記配列を小配列ごとに分散し、記憶した部分には0を挿める。例えば3個の要素からなる配列{3,1,4,1,5,9,2,6,5,3,5,8}であれば、4個の要素のみなる配列{0,0,0,0}を追加して、16の要素からなる配列{3,1,4,1,5,9,2,6,5,3,5,8,0,0,0,0}とする。演算処理装置としては前記8台の演算処理装置のほかに2台の演算処理装置（識別番号8，9とする）を加えた8台の演算処理装置を用いる。この後は、上記第1の実施の形態において詳述した手順により、8台の演算処理装置間で通信を行いデータを交換する。

【0052】本実施の形態においては、一般に（2^ｍ－ｋ）台の演算処理装置に対して、（2^ｍ－ｋ）台の演算処理装置を加えた2^ｍ台の演算処理装置群を構成し、また配列についてもその要素を2^ｍ個に拡張して各演算処理装置に分配し、上記第1の実施の形態と同様の方法で並列計算及びデータの集結を行うものとする。これにより、2の冪乗ではない台数の演算処理装置に対しても2の冪乗の場合に準じた構成とすることで、上記第1の実施形態と同様の作用効果を得ることができる。

【0053】次に、本発明にかかる並列計算機システムの第4の実施の形態について説明する。第2及び第3の実施の形態は、配列の分割数が2の冪でない場合、すなわち（2^ｍ－ｋ）個に分割される場合について、2^ｍ台の演算処理装置によってデータ配列を1個に集結する方法について述べたものである。これに対し本実施の形態は、配列の分割数が、2^ｍ－2^ｎ（ｍ＞ｎ）である場合に対し、（2ｍ－2ｎ）台の演算処理装置で処理するものである。

【0054】本実施の形態における並列計算機システムの演算処理装置間の通信方法として、ここではまず例として、配列を8分割して8台の演算処理装置（識別番号0，1，…，5）に割り当てている場合について説明する。図8はこの場合の演算処理装置間通信方法を時系列で示すチャートである。

【0055】まず、6台の演算処理装置を2つのグループに分割する。演算処理装置グループ1は識別番号0～3の4台で構成される。演算処理装置グループ2は識別番号4～5の2台で構成される。次に、演算処理装置グループ1の4台間、および演算処理装置グループ2の2台間で、上述の第2の実施の形態における手順により、各々のグループでデータを集結させる。図8における第1及び第2ステップがこれに相当する。

【0056】この後、グループ1とグループ2でデータ交換を次の手順で行う。

(1) 演算処理装置4→演算処理装置0、演算処理装置5→演算処理装置2を同時並列に実施。（図8の第3ステップに相当。）

(2) 演算処理装置1→演算処理装置4，演算処理装置3→演算処理装置5を同時並列に実施。（図8の第4ステップに相当。）

(3) 演算処理装置0→演算処理装置1，演算処理装置2→演算処理装置3を同時並列に実施。（図8の第5ステップに相当。）

この方法により、8台の演算処理装置によってデータ配列を集結させることができる。

【0057】また、本実施の形態のもう一つの例として、配列を16分割して16台の演算処理装置（識別番号0，1，…，15）に割り当てている場合について説明する。図7はこの場合における演算処理装置間通信方法を時系列で示すチャートである。

【0058】まず、16台の演算処理装置を2つのグループに分割する。演算処理装置グループ1は識別番号0，1，…，7の8台で構成される。演算処理装置グループ2は識別番号8，9の2台で構成される。次に演算処理装置グループ1の8台の演算処理装置間、および演算処理装置グループ2の2台の演算処理装置間で、上記第2の実施の形態において述べた方法により、各々のグループでデータを集結させる。図7における第1，第2及び第3ステップがこれに相当する。

【0059】この後、グループ1とグループ2でデータ交換を次の手順で行う。

[1'] 演算処理装置8→演算処理装置0、演算処理装置9→演算処理装置4を同時並列に実施。（図7の第4ステップに相当。）

[2'] 演算処理装置3→演算処理装置5，演算処理装置7→演算処理装置9，演算処理装置0→演算処理装置2，演算処理装置4→演算処理装置6を同時並列に実施。（図7の第5ステップに相当。）

[3'] 演算処理装置0→演算処理装置1，演算処理装置4→演算処理装置5，演算処理装置2→演算処理装置3，演算処理装置6→演算処理装置7を同時並列に実施。（図7の第6ステップに相当。）

【0060】この方法により、5台の演算処理装置によ

ってデータ配列を集結させることができる。なお、グループ2からグループ1に送信されたデータのグループ2内の分配は Binary Treeの方式によっている。

【0084】以下、本発明にかかる並列計算機システムの第5の実施の形態について説明する。本実施の形態における演算処理装置間の通信方法は、上記第5の実施の形態の通信方法を一部変更したものである。以下、例として配列を22分割して22台の演算処理装置（識別番号0，1，…，21）に割り当てている場合について説明する。図3及び図8はこの配列に分割の場合における演算処理装置間通信方法を時系列でかがチャートである。図8において第1ステップから第4ステップまでを示した。

【0085】$22=2^4+2^2+2^1$ であるから、まず、演算処理装置を次の3グループに分ける。

グループ1；　識別番号0，1，…，15の演算処理装置（16台）

グループ2；　識別番号16，17，18，19の演算処理装置（4台）

グループ3；　識別番号20，21の演算処理装置（2台）

【0086】次に、演算処理装置グループ1の16台間、演算処理装置グループ2の4台間、および演算処理装置グループ3の2台間で、上記第1の実施の形態の方法により各々のグループでデータを集結させる。これは図8に示した第1ステップから第4ステップまでが相当する。

【0087】この後は、上記第2或いは第3の実施の形態において説明した方法と同様の手順により、データの＊グループ間交換を行う。以下そのデータの通信方法を順を追って説明する。まず、第3ステップでグループ2において、データの集結が終了するが、その時点で既にグループ3のデータの集結は完了しているから、次のステップとして、グループ2の演算処理装置16，18とグループ3の演算処理装置19，20との間でそれぞれデータの交換が行なわれる。これは図8に示したグループ2とグループ3における第3ステップに相当する。この時点で、グループ3の全ての演算処理装置にはグループ2及びグループ3におけるデータがすべて集結された状態となる。

【0088】次に、グループ2及びグループ3の全てのデータが格納されたグループ2の演算処理装置17，18から、それぞれグループ2の演算処理装置17，15に送りてグループ3より受信したデータが送信される。これは図8に示したグループ2における第4ステップに相当する。

【0089】グループ1においては第4ステップで各演算処理装置間でデータの集結が終了するが、次のステップとして、グループ1と、グループ2、3との間でデータの送受信を行う。まず、グループ1の演算処理装置0，1，2，3，4，5から、それぞれグループ2，3の演算処理装置15，17，18，19，20，21に対してデータが送信される。これによりグループ2，3においてはグループ1，2，3の22台の全ての演算処理装置のデータの集結が完了する。これは図9に示した第5ステップに相当する。

【0090】次に、グループ1の16台の演算処理装置を4つのグループに分割する、すなわち、

グループ1の小グループ1；　演算処理装置0，1，2，3

小グループ2；　演算処理装置4，5，6，7

小グループ3；　演算処理装置8，9，10，11

小グループ4，　演算処理装置12，13，14，15

とする。

【0091】この各小グループから1台ずつ演算処理装置を選択する。ここでは演算処理装置0，4，8，12を選択する。この4台の演算処理装置に対して、それぞれグループ2の演算処理装置16，17，18，19から、グループ2及びグループ3に関して集結されたデータを送信する、これは図9に示した第6ステップに相当する。

【0092】次に、グループ1の各小グループにおいて、従来の Binary Tree の方式で演算処理装置間でグループ2，3に関するデータの送受信を行なう。小グループの全演算処理装置においてグループ1，2，3のデータを集結させる。例えば小グループ1においては演算処理装置0から演算処理装置2に対してデータを送信し、次に演算処理装置0，2からそれぞれ演算処理装置1，3に対してデータの送信を行う。他の小グループにおいても同様である。これは図9に示した第7ステップ及び第8ステップに相当する。こうして、全ての22台の演算処理装置において各個のデータ配列の集結を完了する。

【0070】一般に、この事例では表されない台数の演算処理装置におけるデータ配列は、以上説明した方法によって集結させることができる。まず、k個の整数 $n_1, n_2, n_3, \cdots, n_k$（但し、$n_1 > n_2 > n_3 > \cdots > n_k \geq 0$）

を用いて、並列計算機システムの演算処理装置の台数を

【0071】

【数14】

$$2^{n_1}+2^{n_2}+2^{n_3}+\cdots+2^{n_k}$$

と表す。また、データ配列をこの台数と同数の小配列に分割し、各演算処理装置に分割して演算処理を行なうものとする。並列計算機システムの演算処理装置のうち、

【0072】

【数15】

$$2^{n_1}, 2^{n_2}, 2^{n_3}, \cdots, 2^{n_k}$$

台をそれぞれグループ $G_1, G_2, \cdots, G_k$ として、並列計算機システムの演算処理装置を k 個のグループに分

割する。同様にデータ配列の分割列の
【００７３】
【数１６】

$$2^{n_1}, 2^{n_2}, 2^{n_3}, \cdots, 2^{n_k}$$

個をそれぞれ配列Ａ$_1$，Ａ$_2$，…，Ａ$_k$としてｋ個の配列に分割する。
【００７４】次に、１≦ｐ≦ｋなるすべてのｐに対して、以下の『』内に定義する操作（以下、グループ内工程ｐという。）を行う。但し、グループ内工程１，…，ｋは並列して行うこととする。
【００７５】『　グループＧ$_0$の（２のn_1乗）個の演算処理装置に識別番号０，１，…，（２のn_1乗－１）を付与する。次に、０≦ｑ≦ｐ－１なるｑに対し、以下の《　》内に定義する操作ｑを、ｑ＝０からｑ＝ｐ－１まで順次行なう。
《　識別番号Ｎの演算処理装置に対し、２進法で表した識別番号Ｎの２^{q}の位を反転させた番号Ｎ'を識別番号とする演算処理装置を対応させ、データ配列の演算処理結果を識別番号Ｎの演算処理装置と識別番号Ｎ'の演算処理装置との間で相互に送受信する。但し、ｑ＞０なるｑに対しては、操作ｑの際に、識別番号Ｎ、Ｎ'の演算処理装置間で、各演算処理装置による演算処理結果に加えて操作（ｑ－１）までで得られた演算処理結果を合わせて送受信することとする。　》
この操作により、グループＧ$_0$の（２のn_1乗）台の演算処理装置で、データ配列の集結を行う。　』
グループの設定方法により、グループ内工程１，…，ｋを並列に行なったとき、グループ内工程ｋが最初に終了し、以下、グループ内工程（ｋ－１），…，２，１の順に終了する。このことを考慮して、以下の『』に定義する操作（以下、グループ内工程ｐという。）を、ｐ＝ｋ－１からｐ＝１までｐに関して降順に行うこととする。
【００７６】『グループ内工程ｐが終了した後、グループＧ$_0$の各演算処理装置に集結された配列Ａ$_p$のデータを、グループＧ$_0$の演算処理装置から、グループＧ$_{n_1}$，…，Ｇ$_{n_1+\cdots+n_{p-1}}$に属する全ての演算処理装置に送信する。すなわち、グループＧ$_0$に属する（２のn_p乗）台の演算処理装置のうち
【００７７】
【数１７】

$$2^{n_p+1}+\cdots+2^{n_k}$$

台を選択して、これら選択された演算処理装置とグループＧ$_{n_1}$，…，Ｇ$_{n_1+\cdots+n_{p-1}}$に属する演算処理装置とを１対１に対応させ、グループＧ$_0$からグループＧ$_{n_1}$，…，Ｇ$_{n_1+\cdots+n_{p-1}}$へのデータ通信を行う。次に、グループＧ$_{n_1}$からグループＧ$_0$へのデータの送信を行う。（２のn_p乗）台の演算処理装置からなるグループＧ$_0$を、それぞれが
【００７８】
【数１８】

$$2^{n_p-n_{p+1}}$$

台の演算処理装置からなる小グループα_1，…，α_rに分割する。この小グループの数ｒは、
【００７９】
【数１９】

$$r=2^{n_{p+1}}$$

である。ここで、グループＧ$_{n_1}$に属する演算処理装置をｂ$_1$，…，ｂ$_r$と表記する。　グループＧ$_0$の小グループα_1，…，α_rと、グループＧ$_{n_1}$に属する演算処理装置をｂ$_1$，…，ｂ$_r$とを１対１に対応させて、グループＧ$_{n_1}$の演算処理装置ｂ$_i$から対応する小グループα_iのうちから選択された１台の演算処理装置ａ$_i$に、グループＧ$_{n_1}$において集結された配列Ａ$_{n_1}$のデータを送信する操作を、１≦ｉ≦ｒなる全てのｉについて並列に行う。このとき、ｐ＜ｋ－１の場合、演算処理装置ｂ$_1$からａ$_1$へは、グループＧ$_{n_1}$，…，Ｇ$_{n_1+n_2}$より受信したデータ配列Ａ$_{n_1}$，…，Ａ$_{n_1+n_2}$を含めて送信するものとする。
【００８０】この後、各小グループα_iにおいて、演算処理装置ａ$_i$からａ$_1$以外の全ての演算処理装置に対して、従来のBinary Tree の方式でデータの送信を行なう。これにより、グループＧ$_0$の全ての演算処理装置に対してグループ配列Ａ$_1$，…，Ａ$_k$に関するデータ配列の集結が完了する。　』
この方法により、一般に複数台の演算処理装置によって各演算処理装置において分散され並列計算されたデータ配列を、効率よく集結させることができるから、計算の高速化を図ることができる。
【００８１】
【発明の効果】以上説明したように本発明によれば、並列計算機システムの演算処理装置間の通信方法の効率をより向上させることにより、データの授受の際の待ち時間を最小限に抑えることができるから、並列計算機システムにおいて実施される大規模な計算の高速化を図ることができる。
【図面の簡単な説明】
【図１】本発明の第１の実施の形態における並列計算機システムの構成を示すブロック図である。
【図２】本発明の第１の実施の形態にかかる並列計算機システムの演算処理装置間の通信方法を時系列で示すチャートである。
【図３】通信されるデータ量が少ない場合の本発明の第１の実施形態及び従来の通信方法を用いた場合の演算処理台数と通信時間の相関を示すグラフである。
【図４】通信されるデータ量が多い場合の本発明の第

の実施形態及び従来の通信方法を用いた場合の演算処理台数と通信時間の相関を示すグラフである。

【図６】本発明の第２の実施の形態にかかる並列計算機システムの演算処理装置間の通信方法を時系列で示すチャートである。

【図８】本発明の第２の実施の形態にかかる並列計算機システムの演算処理装置間の通信方法を時系列で示すチャートである。

【図７】本発明の第４の実施の形態にかかる並列計算機システムの演算処理装置間の通信方法を時系列で示すチャートである。

【図８】本発明の第５の実施の形態にかかる並列計算機システムの演算処理装置間の通信方法を時系列で示すチャートである。

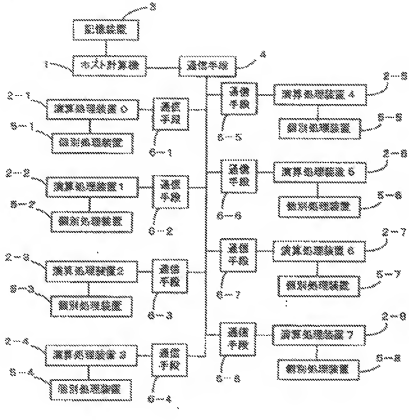
【図９】本発明の第５の実施の形態にかかる並列計算機システムの演算処理装置間の通信方法を時系列で示すチャートである。

【符号の説明】

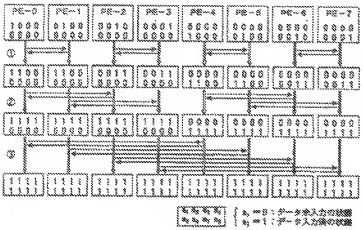
１…ホスト計算機、　２－１…演算処理装置、　３…記憶装置、４…通信手段、　５－１…個別処理装置、　６－１…通信手段、

Ｉｄａ、Ｉｄｂ…本発明の第１の実施の形態における演算処理装置の台数と通信に要する時間の関係を示す曲線

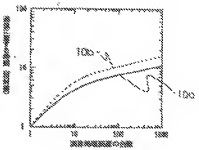
【図１】



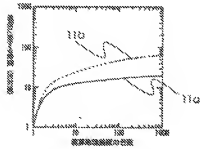
【図２】


PE-0
PE-1
PE-2
PE-3
PE-4
PE-5
PE-6
PE-7
①
②
③

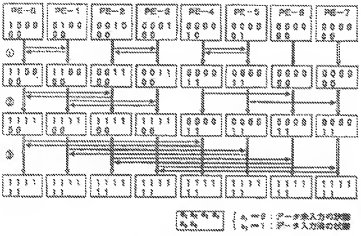
【図３】


100
10
1
10
100
1000

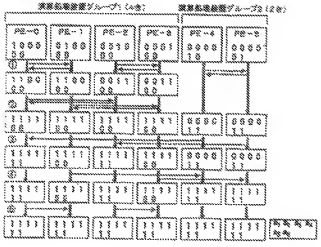
【図４】


1000
100
10
1
10
100
1000

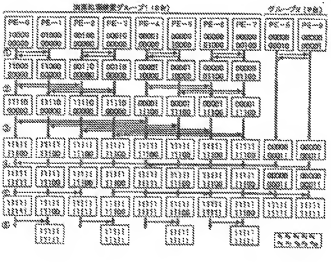
【図５】


PE-0
PE-1
PE-2
PE-3
PE-4
PE-5
PE-6
PE-7
①
②
③

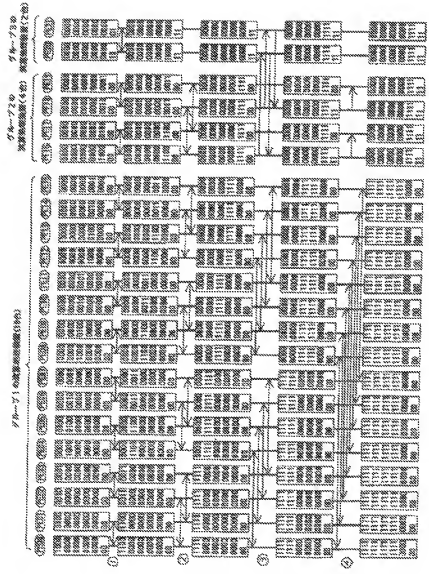
【図6】



【図7】



[図8]



[図6]